

SALÃO DE
INICIAÇÃO CIENTÍFICA
XXIX SIC

UFRGS
PROPESQ



múltipla 
UNIVERSIDADE
inovadora  inspiradora

Evento	Salão UFRGS 2017: SIC - XXIX SALÃO DE INICIAÇÃO CIENTÍFICA DA UFRGS
Ano	2017
Local	Campus do Vale
Título	BigHybrid: Um Toolkit Para a Simulação de Aplicações Big Data Sobre Infraestruturas de Nuvem e Computação Voluntária
Autor	VINÍCIUS PITTIGLIANI PEREGO
Orientador	CLAUDIO FERNANDO RESIN GEYER

BigHybrid: Um Toolkit Para a Simulação de Aplicações Big Data Sobre Infraestruturas de Nuvem e Computação Voluntária.

Nome: Vinícius Pittigliani Perego

Orientador: Cláudio Fernando Resin Geyer

Projeto: SMART-SENT Desenvolvimento de Plataforma para Análise de Big Data em Aplicações de Cidades Inteligentes

Instituição: UFRGS - Grupo de Processamento Paralelo e Distribuído

A contínua instrumentação da sociedade e o fácil acesso à Internet levaram ao intenso uso das mídias sociais, dispositivos móveis, redes de sensores e comércio eletrônico - a era da Internet das Coisas (IoT). Como consequência, a todo momento são enviados milhões de e-mails, mensagens em redes sociais, fotos, vídeos, arquivos de dados e logs de sensores - propriedades de Big Data. Desse modo, é possível interligar pessoas, dispositivos móveis e sensores (IoT) possibilitando a observação, coleta, análise e disseminação de informações (Big Data). Essas características participam de um cenário conhecido como Cidades Inteligentes (CI).

CI necessitam utilizar ferramentas abrangentes para lidar com o gerenciamento e processamento de informações, afim de apoiar aplicações em contextos variados. Em resposta a isso, a Plataforma Smart-Senti (PSS) foi criada e, entre seus principais objetivos pode-se encontrar: (i) O desenvolvimento de estratégias para coleta, processamento e armazenamento de dados em larga escala; (ii) O desenvolvimento de estratégias para o escalonamento de aplicações Big Data em ambientes de multi-cloud; (III) O desenvolvimento de soluções de infraestrutura ou de algoritmos para Cloud, Edge Computing (EDGE) e IoT. EDGE é um método para otimizar recursos em infraestruturas de Cloud que consiste no processamento dos dados em sua origem.

Desse modo, este trabalho apresenta o desenvolvimento de uma solução para a análise e processamento de dados em larga escala, denominada *BigHybrid* (BH). O BH é um *toolkit* de simulação implementado sobre a ferramenta *SimGrid*. O *toolkit* BH permite avaliar diferentes estratégias de processamento sobre ambientes de computação voluntária e de *nuvens*. Os resultados obtidos do simulador apresentam boa acurácia, com aproximadamente 5% de erro em relação ao cenário real. Além disso, a partir dos resultados obtidos, foi possível propor o uso de estratégias para a distribuição de dados e o escalonamento de tarefas de aplicações Big Data processadas em ambientes híbridos. Como consequência das estratégias propostas, o ambiente de processamento pode ser aperfeiçoado de tal forma que o uso de recursos possa ser alocado adequadamente ao consumo da aplicação.

Palavras chave:

Big Data, MapReduce, Infra-estrutura Híbrida, Sistemas Distribuídos, Computação em Nuvem, Grade de Desktop.